A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT

Abstract

Recently, ChatGPT, along with DALL-E-2 and Codex,has been gaining significant attention from society. As a result, many individuals have become interested in related resources and are seeking to uncover the background and secrets behind its impressive performance. In fact, ChatGPT and other Generative AI (GAI) techniques belong to the category of Artificial Intelligence Generated Content (AIGC), which involves the creation of digital content, such as images, music, and natural language, through AI models. The goal of AIGC is to make the content creation process more efficient and accessible, allowing for the production of high-quality content at a faster pace. AIGC is achieved by extracting and understanding intent information from instructions provided by human, and generating the content according to its knowledge and the intent information. In recent years, large-scale models have become increasingly important in AIGC as they provide better intent extraction and thus, improved generation results. With the growth of data and the size of the models, the distribution that the model can learn becomes more comprehensive and closer to reality, leading to more realistic and high-quality content generation. This survey provides a comprehensive review on the history of generative models, and basic components, recent advances in AIGC from unimodal interaction and multimodal interaction. From the perspective of unimodality, we introduce the generation tasks and relative models of text and image. From the perspective of multimodality, we introduce the cross-application between the modalities mentioned above. Finally, we discuss the existing open problems and future challenges in AIGC.